Cross-reference to Related Applications

C1 This application claims priority under 35 U.S.C. § 371 to PCT Application PCT/US98/15563 filed on July 28, 1998, which claims priority from U.S. Application No. 60/254,156, filed July 29, 1997, hereby incorporated by reference.

Please replace the paragraph located on page 16 and beginning with "Fig. 11A" with the following paragraph.

Fig. 11A is a graph showing the correlation of T-cell epitopes with sequence conservation in HIV gp120. Epitope identifications were obtained from the HIV Immunology Database, Los Alamos National Labs, which includes a compilation of epitopes published prior to the year 1995 (see Fig. 11B). Sequence positions found in

C2 multiple epitopes are scored multiple times. Epitope scores were smoothed by an 11-residue averaging window. Sequence conservation was evaluated for twenty-three gp120 sequences. The most popular residue at each position was determined, and then scored for its use at that position in each sequence. Conservative substitutions were counted the same as identical matches. Conservation scores were averaged with a 11-residue window. Epitopes are located in regions with above average conservation (>0.87).

Please replace the paragraph located on page 17 and beginning with "Fig. 11B" with the following paragraph.

C3

Fig. 11B is the sequence of HIV gp120 (SEQ ID NO: 1) obtained from the HIV Immunology Database, Los Alamos National Labs showing the 36 epitope identifications of epitopes published prior to the year 1995.

Please replace the paragraph located on page 17 and beginning with "Figs. 13A" with the following paragraph.

C4

Figs. 13A, 13B, and 13C are graphs showing the combination of predictive methods based on MHC binding and preferred processing at poorly conserved regions. In Fig. 13A, the raw experimental epitope scores are shown. In Fig. 13B epitopes were predicted by the EpiMatrix analysis (no geographic bias in MHC preferences, 20% match to motifs allowed). In Fig. 13C, the EpiMatrix scores were set to zero when the sequence conservation at that residue fell below the average for the whole protein. Note in particular the elimination of predicted epitopes in the segment 130-200.

Please replace the paragraph on page 26 and beginning with "As was observed" with the following paragraph.

As was observed for lysozyme, *M. leprae* cpn10, and staphylococcal nuclease, helper T-cell epitopes in *H*uman *I*mmunodeficiency *V*irus (HIV) gp120 tend to cluster near sites that may be preferentially cleaved during antigen processing (Fig. 6). Epitopes were defined using a variety of T-cell stimulation systems, for example, with lymphocytes from draining lymph nodes of mice immunized with native gp120 (Cease *et al.*, Proc. Natl. Acad. Sci. USA 84: 4249-4253, 1987), peripheral blood lymphocytes from humans immunized with vaccinia virus expressing gp120 (Berzofsky *et al.*, Nature 334: 706-708, 1988), and peripheral blood lymphocytes from HIV patients (Clerici *et al.*, Nature 339: 383-385, 1989). Data have been collected and published on the World Wide Web in the HIV Immunology Database, Los Alamos National Labs. Epitopes are broadly distributed over the C-terminal half of gp120. Fewer epitopes occur in the N-terminal half of the protein, although there is a cluster in the region, 101-119, including the "T2" epitope (Cease *et al.*, *supra*). Overlapping epitopes may be grouped into eight regions (shown as gray boxes in Fig. 6) that encompass most of the gp120 sequence: 31-54, 64-84, 101-119, 203-269, 273-301, 306-369, 417-453, and 457-502. However, several much shorter segments are over-represented in the sample of reported epitopes. We define immunodominant sequences as those occurring in at least four epitopes: 104-115, 225-236, 294-297, 311-349, 426-440, and 486-500. The number of immunodominant regions (shown as black boxes in Fig. 6) is very similar to the number